ANALYSIS AND APPLICATION OF MINIMUM VARIANCE DISCRETE TIME SYSTEM IDENTIFICATION*

S. Kotob
Computer Sciences Corporation
6565 Arlington Blvd.
Falls Church, VA 22041

H. Kaufman
Department of Electrical and
Systems Engineering
Rensselaer Polytechnic Institute
Troy, NY 12181

Abstract

An on-line minimum variance parameter identifier is developed which embodies both accuracy and computational efficiency. The formulation results in a linear estimation problem with both additive and multiplicative noise. The resulting filter which utilizes both the covariance of the parameter vector itself and the covariance of the error in identification is proven to be mean square convergent and mean square consistent. The MV parameter identification scheme is then used to construct a stable state and parameter estimation algorithm.

1. Introduction

In designing adaptive control systems, it is necessary to determine whether to implement an explicit system in which on-line parameter identifier is needed or an implicit system which does not require explicit parameter identification. Recent studies have indicated preference for explicit designs whenever the process to be controlled has non-minimum phase characteristics and/or high gain and large bandwidth limitations. [1,2]

The development of such an adaptive control system requires the use of an identification scheme that is capable of supplying parameter estimates at an accuracy and rate specified by the controller characteristics. Because a digital adaptive controller uses elements of the discretized matrices, identification of these elements and not the continuous physical system parameters should be considered. Furthermore, identification of the parameters of a continuous system (e.g., stability derivatives) from discrete data results in a problem with many severe nonlinearities.

Linear system identification using the input and noisy measurements of the output can be generally cast as a state estimation problem with both additive and multiplicstion noise (AMN). These terms will in fact be functions of the same noise sequence. The continuous optimal nonlinear filter as derived by Kushner[3] for (AMN) is infinite dimensional and its physical realization is impossible. Approximate linear filters were subsequently derived for AMN[4,5] under the assumption that the additive and multiplication disturbance terms are functions of two independent random processes; hence these results are not immediately applicable to system identification. Thus, a new on-line minimum variance filter for the identification of systems with additive and multiplicative noise has been developed which embodies both accuracy and computational efficiency. The resulting filter is shown to utilize both the covariance of the parameter vector itself and the covariance of the error in identification. A bias reduction scheme can be used if desired, to yield asymptotically unbiased estimates.

As common in deriving any estimation scheme, proof of the convergence of the identification filter is an integral part of the validation of the results. In this respect, the proposed identification scheme is shown to be convergent in the mean square sense. The proof consists of deriving a suitable upper-bound for the mean square error (MSE) and showing that the MSE converges to zero as time tends to infinity. The mean square convergence of the filter implies convergence with probability which, in turn, would imply that the estimates are consistent. Using the proposed parameter identification filter and the related convergence proofs, a state-parameter estimation scheme is constructed and proven to be stable in the sense of boundness. The resulting state-parameter scheme is shown to be computationally feasible and amenable for on-line system identification and adaptive control applications.

To illustrate the reliability of the identification schemes and the problems encountered, experimental results for simulated linearized lateral aircraft motion in a digital closed loop mode, are included. A comparison of the extended Kalman filter and the minimum variance filter in the adaptive mode are presented.

* This work was supported by NASA Grant Number NGR 33-018-183 administered by Langley Research Center, Hampton, VA.

2. Problem Definition

The problem of determining on-line values of certain parameters appearing in the discrete equations of a linear constant coefficient process, which are best with regard to use in adaptive control logic, given the input and noisy measurements of the output, was considered using an on-line minimum variance filter. The corresponding equations are:

$$x(k+1) = A\,x(k) + B\,u(k) \tag{1}$$

$$y(k) = x(k) + \eta(k) \tag{2}$$

where

- $x(k)$ = plant state at the k^{th} sample instant $(n \times 1)$
- A = state transition matrix for the discrete system $(n \times n)$
- $u(k)$ = control vector $(m \times 1)$
- B = control distribution matrix $(n \times m)$
- $y(k)$ = measurement vector $(\ell \times 1)$ at the k^{th} instant
- $\eta(k)$ = measurement noise at the k^{th} instant covariance matrix $R(i,j) = \sigma^2\,\delta_{ij}$, where $\delta_{ij}=1$, $i=j$, $\delta_{ij}= 0$ otherwise.

3. Parameter Modeling

To estimate any unknown vector of parameters q appearing in the state transition matrix A and in the control distribution matrix B, it is necessary to model the dynamics and observations of the system parameters. Furthermore, since not all parameters appearing in the matrices A and B are to be identified, the differentiation between the set of parameters that are to be identified and the set of parameters not to be identified is generally recommended. In particular, a convenient representation of the system is:

$$x(k+1) = C(k)\,q(k) + D(k)\,S \tag{3}$$

where

C and D are selection matrices containing values of the system state and control at the k^{th} instant. A zero entry for a particular C_{ij} (or D_{ij}) would indicate that no coupling exists between x_i and q_j (or between x_i and s_j).

q is a vector of unknown parameters appearing in the A and B matrices.

S is a vector of known parameters appearing in the A and B matrices.

The model for the constant deterministic or stochastic parameter vector q is then given by:

Systems Dynamics

$$q(k+1)=q(k),\ E(q(0))=q_0$$

$$P_0=E[(q(0)-q_0)(q(0)-q_0)^T] \tag{4}$$

Observation

$$x(k+1) = C(k)\,q(k) + D(k)\,S \tag{5}$$

The identification problem as defined in (4) and (5) appears to be a conventional linear state estimation problem. However, because x is not known exactly, C and D are also unknown, and the rules and usage of the conventional Kalman filter cannot be applied. Substituting equation (2) into (5) gives:

$$y(k+1)-\eta(k+1)=C(y(k)-\eta(k),u(k))q(k)$$

$$+ D(y(k)-\eta(k),u(k))\,S \tag{6a}$$

where

the notation $C(y(k)-\eta(k),\ u(k))$ and $D(y(k)-\eta(k),u(k))$ stresses the fact that the selection matrices C and D are functions of the state and control values. Noting that:

$$C(y(k)-\eta(k),u(k))=C(y(k),u(k))-C(\eta(k),u(k))$$

and similarly for D, equation 6a can be rewritten as:

$$y(k+1)-D(y(k),u(k))S=C(y(k),u(k))q(k)$$

$$-C(\eta(k),u(k))q(k)-D(\eta(k),u(k))S+\eta(k+1) \tag{6b}$$

Defining

$$z(k) = y(k+1) - D(y(k),\ u(k))\,S$$

to be a pseudo-measurement vector for the linear system given in (4) and rearranging equation (7) gives the parameter model:

$$q(k+1) = q(k) \tag{7}$$

$$z(k)=\hat{C}(k)q(k)-C_\eta(k)q(k)$$

$$+\eta(k+1)-D_\eta(k)S \tag{8}$$

where

$$\hat{C}(k) = C(y(k),\ u(k))$$

$$C_\eta(k) = C(\eta(k),\ u(k))$$

$$D_\eta(k) = D(\eta(k),\ u(k))$$

Equations (7) and (8) denote a linear time invariant system with a transition matrix I and observation matrix C(k). The observation is corrupted by the multiplicative noise term $C_\eta(k)$ and the additive noise terms $\eta(k)-D_\eta(k)S$ with covariance $R_{eq}=R+E\,\{D_\eta S\,\tilde{S}^T\,D_\eta^{\ T}\}$

4. Minimum Variance Estimation

4.1 Development

Because of divergence and/or inaccuracies common to most existing identification schemes, it was desirable to develop an alternate scheme that could hopefully deal with these problems. The proposed filter is based on a minimum variance performance index for the state estimation of a linear system with additive and multiplicative noise.

The optimum minimum variance filter for a continuous system is in fact nonlinear [3], and its exact implementation is virtually impossible. A linear optimal filter was therefore of interest. Thus defining the identification algorithm to be

$$\hat{q}(k)=\hat{q}(k-\ell)+K(k)[z(k)-\hat{C}(k-1)\hat{q}(k-\ell)] \tag{9}$$

K is to be determined so as to minimize:

$$J = \sum_i E_y(q_i(k) - \hat{q}_i(k))^2 \tag{10}$$

which is the trace of the covariance matrix

$$P(k)=E_y\{(q(k)-\hat{q}(k))(q(k)-\hat{q}(k))^T\} \tag{11}$$

where $E_y(\alpha(k))=E[\alpha|y(0),\ldots y(k)]$

The parameter ℓ (which will be discussed in the next section) defines the frequency of identification. Define the parameter error as $\tilde{q}(k)=\hat{q}(k)-q(k)$, with initial conditions at T=0

$$E(\tilde{q}(0))=0 \ ; \ E\{\tilde{q}(0)\ \tilde{q}(0)^T\} =P_0$$

Using (7), (8) and (9) the error $\tilde{q}$ propagates as:

$$\tilde{q}(k)=\tilde{q}(k-\ell)+K(k)[-\hat{C}(k-1)\tilde{q}(k-\ell) -C_\eta(k-1)q(k)-D_\eta(k)S+\eta(k)] \tag{12}$$

By post multiplying 12 by its transpose, and taking the conditional expectation over the entire measurement vector history $(y(0),\ldots y(k))$ and noting that:

$$E[C_\eta(k-1)q(k)\tilde{q}^T(k-\ell)|y_0,\ldots y(k)]=0$$

when $\ell > 1$, the difference equation for the conditional variance P(k) becomes:

$$P(k)=P(k-\ell)-K(k)\hat{C}(k-1)P(k-\ell)-P^T(k-\ell)\hat{C}^T(k-1)K^T(k) +K(k)(\hat{C}(k-1)P(k-\ell)\hat{C}^T(k-1)+\omega(k-1)+R_{eq})K^T(k) \tag{13a}$$

where

$$\omega_{k-1}=E\{C_\eta(k-1)q(k-\ell)q^T(k-\ell)C_\eta^T(k-1)\} \tag{13b}$$

Stationary conditions for the minimization of the trace of P(k) are obtained by setting all derivatives of (12) with respect to the elements of K(k) equal to zero. This yields:

$$K(k)=P(k-\ell)\hat{C}^T(k-1)[\hat{C}(k-1)P(k-\ell)\hat{C}^T(k-1) +\omega(k-1)+R_{eq}]^{-1} \tag{14}$$

4.2 Observations

(1) The gain of the resulting filter is a function of the error covariance P and the weighted noise covariance ω(13b); where the weighting matrix for ω is the covariance of the identified parameter q.

(2) The derivation of the proposed minimum variance filter is made possibly by not identifying every sample; i.e., $\ell > 1$. For $\ell = 1$, the expected value of many cross terms involving the parameter q, the error $\tilde{q}$ and the noise selection matrix C_η will not vanish; this can be illustrated by noting that:

$$E\{C_\eta(k-1)\tilde{q}(k)q^T(k-1)|y_0,\ldots y(k)\} \neq 0 \tag{15}$$

5. Bias Reduction

Although the linear minimum variance filter as described by equations (9), (14) and (15) was observed to be relatively accurate with respect to other linear schemes, a substantial bias did appear in the parameter estimates, especially in the estimates of insensitive parameters [2]. An investigation was therefore conducted to determine the causes and the means to reduce or eliminate the bias.

By combining equations (15 and (14), the gain can be rewritten as:

$$K(k)=P(k)C^T(k-1)R_\omega^{-1}(k-1) \tag{16a}$$

where

$$R_\omega=R_{eq} + \omega_{k-1} \tag{16b}$$

By substituting (16) in (9) and taking the expectation and the limit as $K \to \infty$, it is found that:

$$E\{\hat{q}(k)\}=[E\{\hat{C}^T(k-1)R_\omega^{-1}\ \hat{C}(k-1)\}+E\{C_\eta^T(k-1)R_\omega^{-1} C_\eta(k-1)\}]^{-1} \cdot [E\{\hat{C}^T(k-1)R_\omega^{-1}\hat{C}(k-1)\}]\cdot q(k) \tag{17}$$

Obviously equation (17) reveals the bias in the estimates of q, $\hat{q}$. Assuming that the term $E\{\hat{C}_\eta^T(k-1)R_\omega^{-1}\hat{C}(k-1)\}$ is a generalized measure of the signal power, and $E\{C_\eta^T(k-1)R_\omega^{-1}C_\eta(k-1)$ is a generalized measure of the noise, equation (17) can be written as:

$$E\{\hat{q}(k)\}=[\underline{S} + \underline{N}]^{-1}\ [\underline{S}]\cdot q(k) \tag{18}$$

where

$$\underline{S} = E\{C^T(k-1)R_\omega^{-1}\ C(k-1)\}$$

$$\underline{N} = E\{C_\eta^T(k-1)R_\omega^{-1}\ C_\eta(k-1)\}$$

By examining equation (17) and (18), it becomes obvious that the troublesome term is the noise power $\underline{N}$. Hence a correction term must be added so as to compensate for the bias. From consideration of equations (16) and (17), it is clear that the correction term must incorporate the covariance term P_k and the latest estimate $\hat{q}_{k-\ell}$. Adding the correction term to (14), the basic algorithm becomes:

$$\hat{q}(k)=\{I+P(k)\ G(k)\}\hat{q}(k-\ell) + K(k)\ [z(k)-\hat{C}(k-1)\ \hat{q}(k-\ell)] \tag{19}$$

where G(k) is to be found such that

$$\lim_{k\to\infty}\ E\{\hat{q}(k)\} =q \tag{20}$$

Taking the expectation for (19) and using (20), yields:

$$G(k)=E\{C_\eta^T(k-1)R_\omega^{-1}C_\eta(k-1)\}=N$$

Hence, the modified minimum variance filter is given by:

$$\hat{q}(k)=(I+P(k)E\{C_\eta^T(k-1)R_\omega^{-1}\ C_\eta(k-1)\})\hat{q}(k-\ell)$$
$$+\ K(k)[z(k)-\hat{C}(k-1)\hat{q}(k-\ell)] \qquad (21)$$

where $P(k)$ and $K(k)$ are given by the recursive equ equations (14) and (15). It should be pointed out that in recursive on-line parameter identification schemes, only asymptotic unbiasness is possible[6].

6. Filter Stability

Essential to any estimation scheme is the validity of the resulting estimates. In this respect, it is desired to prove that the proposed identification algorithm converges to the actual system parameters. The convergence of the filter is of particular importance since the resulting estimates are to be used in the construction of an adaptive controller. Before proceeding in establishing the convergence of the proposed identification scheme, the following assumptions needed for the proof are stated:

A1. $[\eta_k]$ is a vector sequence whose entries are zero mean independent variables. All entries of the measurement noise vector (η_k) are mutually independent. Second and fourth moments of $[\eta_k]$ are uniformly bounded.

A2. The deterministic control vector $u(k)$ is assumed to be bounded. Similarly the output vector $y(k)$ and all transformations on $y(k)$ are assumed to have bounded moments.

A3. The linear system is completely controllable and completely observable.

A4. The parameter set to be identified is assumed to be completely observable [7] in the sense that the information matrix

$$\mathcal{F}(k,1)\ \underline{\Delta}\ \sum_{i=1}^{k}\ \hat{C}^T(i)R_\omega^{-1}\hat{C}(i)$$

is positive definite.

A5. The product of the matrix P and the signal power S is positive definite.

6.1 Parameter Filter Convergence

Theorem 1. Mean Square Convergence. Under the assumptions A1 to A5, the linear estimator of $q(k)$ given in equations 14, 15 and 21 converges in the mean square sense to the unknown parameter vector q of the linear system in (1-2).

Proof: The estimation error $\tilde{q}$ can be given by the following equation:

$$\tilde{q}(k)=[I+P\ N-P\hat{C}^T(1-1)R_\omega^{-1}\ \hat{C}(k-1)]\tilde{q}(k-\rho)$$
$$+P[\hat{C}^T(k-1)R_\omega^{-1}\hat{C}(k-1)q-Nq$$
$$-\hat{C}^T(k-1)R_\omega^{-1}\hat{C}(k-1)q-\hat{C}^T(k-1)R_\omega^{-1}(\eta-D_\eta S)] \qquad (22)$$

Premultiplying (22) by $\tilde{q}(k)$, taking the expectation and by repeated use of the Cauchy-Shwartz and triangle inequalities, an upper bound for the mean square of the identification error can be established. [8] By applying Venter's Theorem [9] and using the fact that in the limit, the maximum eigenvalue of P behaves as $\frac{1}{k}$, the mean square error $E[||\tilde{q}(k)||^2]$ is shown to converge to zero in the limit, i.e.,

$$\lim_{k\to\infty} E[||\tilde{q}(k)||^2] = 0$$

Combining the fact that the proposed identification scheme is asymptotically unbiased and mean square convergent, it is concluded that the filter is mean-square consistent.

6.2 Convergence of State Estimation

The maximum likelihood, minimum variance and least squares estimate of the state vector $x(k)$ given the measurement vector $y(0)....y(k)$ is given by the Kalman-Bucy filter. The Kalman filter was shown to converge in the mean square sense and with probability 1 if the plant model and Gaussian noise statistics are exactly known. In cases where the plant model is not exactly known, an approximate Kalman filter can be constructed using identified parameters. The stability of the approximate Kalman filter is discussed in the sequel.

Theorem 2: Given the approximate Kalman filter

$$\hat{x}(k/k)=\hat{A}\ \hat{x}(k-1/k-1)+\hat{B}\ u(k-1)$$
$$+\hat{K}_s(k)[y(k)-[\hat{A}\ \hat{x}(k-1/k-1)+\hat{B}\ u(k-1)]]$$
$$K_s(k)=P_s(k)R_k^{-1} \qquad (23)$$

where $\hat{x}(k/k)\underline{\Delta}$ state estimate using identified parameters

$\hat{A}$, $\hat{B}$ identified system and input matrix

$\hat{K}_s$, $\hat{P}_s$ gain and covariance matrix using identified parameters,

if the linear system (1,2) is stable and if $\hat{A}$ and $\hat{B}$ are consistent estimates of A and B respectively, then

$$\lim_{n\to\infty}\frac{1}{n}\ \sum|x(k|k)-\hat{x}(k|k)|=0 \quad \text{with probability one}$$

$$\hat{P}_s(k) \to P_s^o \quad \text{as } k \to \infty \text{ with probability one}$$

$$\hat{K}_s(k) \to K_s^o \quad \text{as } k \to \infty \text{ with probability one}$$

where

$x(k|k) \triangleq$ optimal state estimate

K_s^o , P_s^o = optimal gain and covariance matrix

The proof of theorem 2 is provided in reference (8).

It was shown in section 6.1 that the estimator $\hat{q}$ in (21) is mean square convergent and mean square consistent. By invoking theorem 2 and the stability condition on the plant in 1, it can be observed that the proposed identifier and the respective approximate Kalman filter constitute a stable (in the sense of boundedness) state-parameter identification scheme. This structure can serve as an alternative to the linearized Kalman filter with the advantages of stability and ease of implementation. Simulations for the state parameter filter are presented in section 8.

7. Identification of Time Varying Parameters

Modeling time varying parameters as a first order random walk, the minimum variance filter can be modified or rederived so as to track the variations in the system parameter $q(k)$. The new parameter model is then given by the following equation:

Parameter Model

$$q(k) = q(k-\ell) + v(k-\ell) \qquad (24)$$

where v is a zero mean uncorrelated stationary Gaussian noise sequence with covariance $Q(k) = E\{v(k)\, v(k)^T\}$.

Since the parameter vector q is modeled as a first order random walk, its covariance has to be updated recursively so as to compute $\omega(k-1)$. Assuming that the initial parameter covariance T is given by:

$$T(0) = E\{q(0)\, q^T(0)\};$$

the parameter covariance $T(k)$ and the weighted noise covariance are given by [8]:

$$T(k) = T(k-\ell) + Q(k-\ell) \qquad (25)$$

$$\omega(k-1)=E\{C_n(k-1)T(k-\ell)C_n^T(k-\ell) \qquad (26)$$

Equations (25), (26) summarize the minimum variance filter for identification of varying parameters. The resulting filter is relatively simple for use in a typical process control computer.

8. Applications and Results

The performance of the minimum variance filter was evaluated experimentally using an adaptive controller designed for the linearized lateral motion of a typical fighter aircraft[10].

In particular feedforward gains were updated by direct formula evaluation, while a Riccati type iterative procedure was used to update the feedback gains.

For evaluation purposes, the aircraft was assumed to be flying in a fixed flight condition (FC2, Mach 0.9, 3000 m). A sensitivity study defined, for identification, a set of 12 parameters which make up the first and third rows of A and B matrices. Parameter estimates were obtained every other sample (ℓ=2) using noisy measurements of the states [2] and used every 1 sec. in the gain adaptation procedure. The resulting parameters and gains were then used to estimate the states and controls each sample period of 0.2 sec. The square wave aileron pilot input u_m of 5° at the frequency of 0.4 Hz was used in all the experiments.

The convergence properties, adaptive controller results and comparisons with different identification procedures to be presented in this paper were all conducted for FC2 and all parameter estimates were initialized at 50% of their actual values.

Figures 1.a,b, and c illustrate the behavior using the minimum variance filter. The asymptotic unbiasness is evident in figures 1.a and c which show that the estimates have converged within 20 sec. (50 measurements). The parameter a_{13} featured in figure 1.b is very insensitive especially with aileron excitation (the parameter a_{13} couples the sideslip angle to the roll rate).

Figure 2 depicts the roll rate behavior in the adaptive mode. It can be seen that model following performance is highly correlated with convergence of the parameter estimates. Reasonable model following was achieved after 15 sec. when most parameters had converged to the actual values. Comparative results using the ELF are shown in Fig. 3.

9. Discussions and Conclusions

A linear minimum variance parameter identifier was derived and was shown experimentally to converge to the actual parameters. A bias reduction scheme and modifications for time varying parameters were presented. The new filter proved superior over existing linear and linearized parameter filters and generally more flexible and effective in the estimation of insensitive parameters.

References

1. Adaptive Control and Guidance for Tactical Missiles, TR 170-1, JASC, 1970.
2. Kaufman, et. al., TASA CR-2466, Dec. 1974.
3. Kushner, J., "Approximations to Optimal Nonlinear Filters", IEEE,G-AC, vol. 12, No. 5.
4. McLane, P.M., "Optimal Linear Filtering for Linear Systems with State-Dependent Noise", Int. J. Control, 1969, vol. 10, No. 1, 41-51.
5. Rajasekaran, P.K., et. al., "Optimal Linear Estimation of Stochastic Signals:, IEEE AES-7, No. 3, May 1971, pp. 462-468.

6. Mendel, J., "Discrete Techniques of Parameter Estimation", Marcel Dekker, Inc., N.Y., 1973.

7. Jazwinski, A.H., Stochastic Processes and Filtering Theory, Academic Press, N.Y., 1970.

8. Kotob, S., "Discrete On-Line System Identification", Ph.D. Thesis, RPI, Troy, NY, 1975.

9. Venter, J.H., "An Extension of the Robbins-Monroe Procedure", Annals Math. State., 38, (1967).

10. Alag, G. and Kaufman, H., "Digital Adaptive Model Following Flight Control", AIAA Mechanics Conf., Calif., August, 1974.

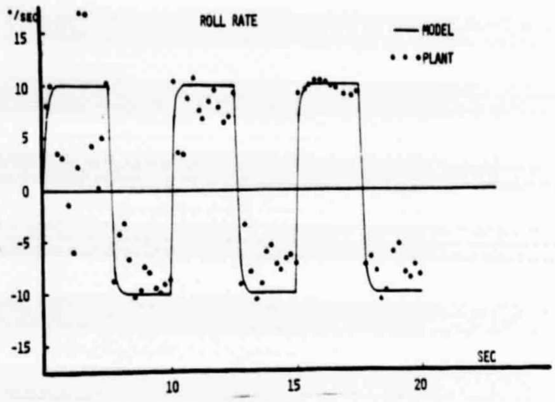

Fig. 2 Roll Rate

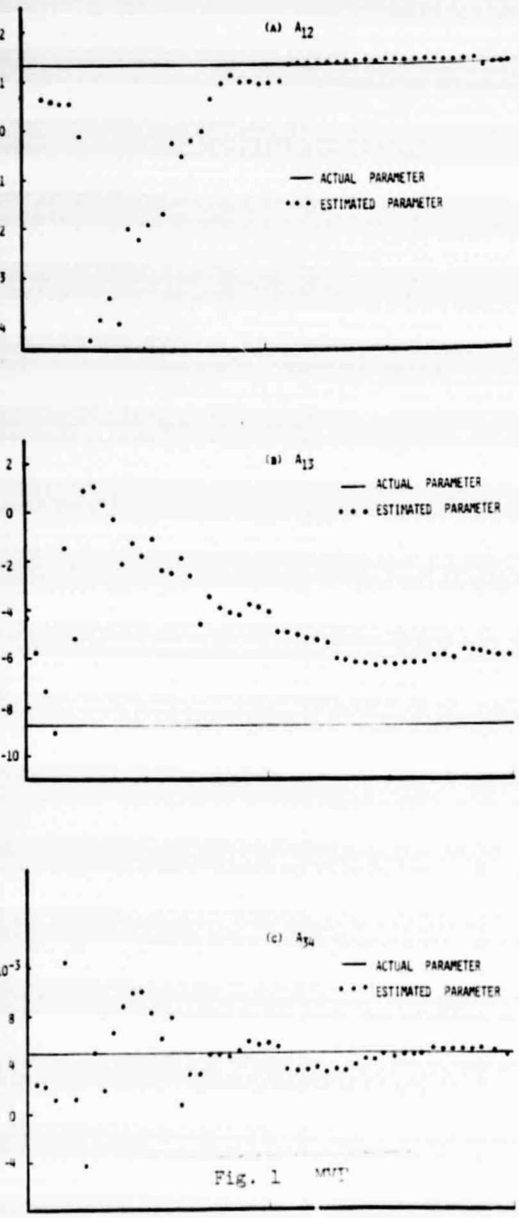

Fig. 1 MVT

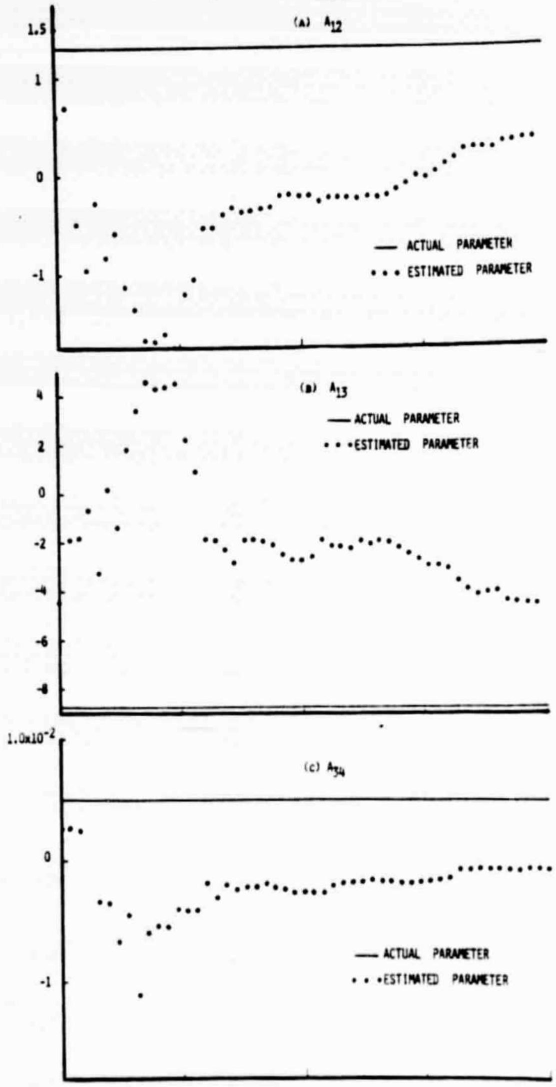

Fig. 3 EKF